AD_____

Award Number: W81XWH-12-1-0179

TITLE: Unbiased Combinatorial Genomic Approaches to Identify Alternative Therapeutic Targets within the TSC Signaling Network

PRINCIPAL INVESTIGATOR: Norbert Perrimon

CONTRACTING ORGANIZATION: Harvard College, President and Fellows of Sponsored Programs Admin
Boston, MA 02115

REPORT DATE: June 2014

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
June 2014	Annual	1 June 2013 – 31 May 2014

4. TITLE AND SUBTITLE

Unbiased Combinatorial Genomic Approaches to Identify Alternative Therapeutic Targets within the TSC Signaling Network

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-12-1-0179

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Dr. Norbert Perrimon and Dr. Brendan Manning

E-Mail: perrimon@receptor.med.harvard.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Harvard College, President and Fellows of
Sponsored Programs Admin
Boston, MA 02115-6027

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The long-term goal of this project is geared toward further defining the regulatory mechanisms impinging on the TSC-Rheb circuit and revealing therapeutic strategies to target this signaling network in genetic tumor syndromes and cancer. Specifically, we aim to establish a robust synthetic lethal screening method applied to the study of the TSC network, validate synthetic lethal pairs in an in vivo intestinal stem cell system, and determine whether synthetic lethal combinations found in Drosophila are relevant to mammalian networks. In the second year of funding we have made significant progress and generated the required mutant cell lines and performed a combinatorial screen that has led to some exciting leads that we are in the process of being validated in mammalian cells.

15. SUBJECT TERMS
TSC, RNAi, synthetic screen, Drosophila

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
U	U	U	UU	52	USAMRMC

Table of Contents

1. INTRODUCTION:

A detailed understanding of how common oncogenic signaling pathways are assembled into larger signaling networks is essential to developing therapeutic strategies to properly target these pathways in cancer and for interpreting clinical outcomes from targeted therapeutics. While the effected oncogenes and tumor suppressors that predominate different classes of human cancer can vary greatly, a small number of highly integrated signaling nodes are affected in the majority of human cancers, regardless of tissue of origin. It is therefore important to understand how these key signaling nodes are regulated. In this project, we focus on one such node, involving the TSC1-TSC2 complex and the Ras related small G protein Rheb, which is aberrantly regulated in nearly all genetic tumor syndromes and the most common forms of sporadic cancer. The long-term goal of this project is geared toward further defining the regulatory mechanisms impinging on the TSC-Rheb circuit and revealing therapeutic strategies to target this signaling network in genetic tumor syndromes and cancer. For this purpose, we will use high-throughput technologies in *Drosophila* to identify synthetic lethal interactions between TSC network tumor suppressors and identified pathway interactors. We will then go on to validate positive hits in an *in vivo Drosophila* model before determining which interactions are conserved using mammalian cell culture.

2. KEYWORDS:

Drosophila, TSC, Drug Targets, Combinatorial Screen, Cancer

3. OVERALL PROJECT SUMMARY:

We have made progress towards all the initial goals. Details are provided below according to the original Statement of Work:

Task 1. Establish a robust synthetic lethal screening method applied to the study of the TSC network. (months 1-36).

- Characterize an optimized shRNA targeting each of the core tumor suppressors. (months 1-2) COMPLETED
- Characterize the tumor suppressor genomic rescue constructs. (months 1-2) COMPLETED
- Clone shRNAs targeting each of the candidate genes into each of these constructs to create the desired pairwise combinations. (months 3-8) COMPLETED
- Establish tumor suppressor cell lines (months 12-24) COMPLETED
- Perform the synthetic screen for viability with kinase and phosphatase set (months 8-14) COMPLETED
- Confirm the positives (months 14-24) COMPLETED
- Perform more quantitative screens using phospho-AKT, dpERK, and phospho-S6K antibodies (months 24-36) UNDERWAY

-Expand the synthetic lethal screen to the Insulin network (months 24-36) UNDERWAY

We have made considerable progress towards performing combinatorial screens for the TSC/Insulin network in *Drosophila* cultured cells. Specifically, we combined the CRISPR genome editing system with a novel approach allowing efficient single cell cloning of *Drosophila* cells with the aim of generating knockout cell lines for five tumor suppressors (Tsc1, Tsc2/gig, Nf1, Lkb1 and Pten) within the TSC/Insulin signaling pathway (Housden et al., 2014).

To enable the use of the CRISPR system to generate mutant cell lines, we first assessed the specificity of mutation in *Drosophila* S2R+ cells. We generated a quantitative mutation reporter vector in which an sgRNA target sequence was cloned into the coding sequence of the luciferase gene immediately following the start codon. A fixed proportion of indels induced by the CRISPR system at this site therefore lead to frame shift of the luciferase gene and ablation of functional protein. 75 sgRNA expressing plasmids were generated with varying number and position of mismatches to the target sequence and cotransfected into S2R+ cells with the reporter construct and luciferase levels measured after 4 days to determine mutation rate. From this analysis, we found that 3bp of mismatch is sufficient to prevent detectible mutation as long as at least 1 mismatch is within the 15bp 3' seed region (Figure 1A). Importantly, this result demonstrates high specificity compared to mammalian systems where mutations have been detected with 5bp of mismatch.

Next, we tested whether sgRNA efficiency could be predicted based on sequence. Previous experiments have demonstrated widely varying mutation rates between sgRNAs so a method of predicting efficiency would greatly facilitate the use of CRISPR to generate mutant cell lines. To achieve this, we generated luciferase reporters similar to that described above, containing 75 different target sequences and corresponding sgRNA expressing plasmids for each. Previous reports have suggested that high GC content at the 3' end of the sgRNA is associated with high efficiency yet our results contradict this with no evident correlation. We analyzed the base pair composition of each position amongst high and low efficiency sgRNAs from the 75 that we tested. From this, we were able to generate a matrix representing the association of each base in each position within the sgRNA with high mutation rate (Figure 1C) and generate a

sequence-based scoring algorithm to predict efficiency. Comparison with independent data from two *Drosophila* publications showed good correlation between the scores we generated and the reported mutation rates (Figure 1D).

Although CRISPR works with high efficiency in cultured *Drosophila* cells, previous attempts to generate mutant *Drosophila* cell lines using CRISPR have failed due to rapid selection for wildtype cells and reversion of the population. We therefore tested whether it would be possible to isolate single mutant cells and culture these to produce cultures completely lacking wildtype sequence at the target locus. No robust methods existed to clone single *Drosophila* cells so we tested an approach where single cells isolated using flow cytometry were seeded into different media formulations. We found that the use of preconditioned media allowed robust survival and expansion of single S2R+ cells in culture (Figure 2A).

Next, we tested whether a CRISPR-treated population contained single cells lacking wildtype sequence (homozygous mutations). 30 single cells were tested using high-resolution melt assays (HRMA) and 21 were identified as carrying mutations. 8 of these were sequenced and wild type sequences were absent from all of them (Figure 2B and data not shown), indicating that the generation of mutant cell lines using this method was likely to be efficient (Figure 2C).

We have successfully produced mutant lines for three of the tumor suppressor genes (Nf1, Tsc1 and gig/Tsc2) and are in the process of generating a mutant line for Pten. We also attempted to generate a mutant line for Lkb1 although severely reduced proliferation rates prevented the maintenance of these cells.

Characterization of the mutant cell lines demonstrated that the observed morphological and cell growth phenotypes are consistent with previously demonstrated in vivo phenotypes (Figure 3). To gain further insight into TSC/Insulin pathway function, we performed phosphoproteome analysis of Tsc1 and gig/Tsc2 mutant lines and results are currently being analyzed. Furthermore, we performed combinatorial screens by treating the mutant lines with dsRNAs targeting all kinases and phosphatases (563 genes) in the *Drosophila* genome. 65 samples that displayed synthetic lethality (15 genes) or synthetic increases in viability (50 genes) with Tsc1 and/or gig/Tsc2 mutations were selected as hits (Figure 4), three of which scored as synthetic lethal with both Tsc1 and gig/tsc2 mutations (mRNA-Cap, Pitslre and CycT). Interestingly, mRNA-Cap is required for the addition of 7-methylguanosine caps to mRNA molecules. This cap is the target of 4EBP regulation downstream of insulin signaling, indicating a likely mechanism for the observed synthetic lethality. We are now in the process of testing whether the homologs of the three hits display similar synthetic lethal interactions in mouse and human cell lines.

In addition, we are preparing to complete a second screen to test the tumor suppressor mutants in combination with RNAi knockdown of the TSC/Insulin pathway network that we have generated. Previously, we have built a high confidence protein-protein interaction (PPI) network for insulin signaling that we interrogated functionally by RNAi using MAPK/ERK (Friedman et al., 2011) and AKT (Kulkarni et al., in preparation) phosphorylation as read outs. To further examine the interactions between components of this network and in particular synthetic genetic interactions with TSC we will test potential synthetic lethal combination in the tumor suppressor cell lines described above.

Task 2: Validation of synthetic lethal pairs in an *in vivo* intestinal stem cell system. (months 1-36).

- Characterize the phenotypes and level of knockdown of single shRNAs targeting the five tumor suppressors in ISCs. (months 1-6) COMPLETED
- Characterize shRNAs targeting each gene that show synthetic phenotypes in combinations with

the tumor suppressors. (months 6-24) COMPLETED
- Characterize in details the phenotypes of the synthetic interactions using phosphoHistone H3, caspase antibodies, and BrDU. (months 12-36) UNDERWAY
- Characterize in further details the phenotypes of the synthetic interactions using pathway specific phospho-antibodies such as phospho-AKT, dpERK, and phospho-S6K (months 14-36) UNDERWAY
- Confirm by genomic rescue the specificity of the interactions (months 15-36). UNDERWAY

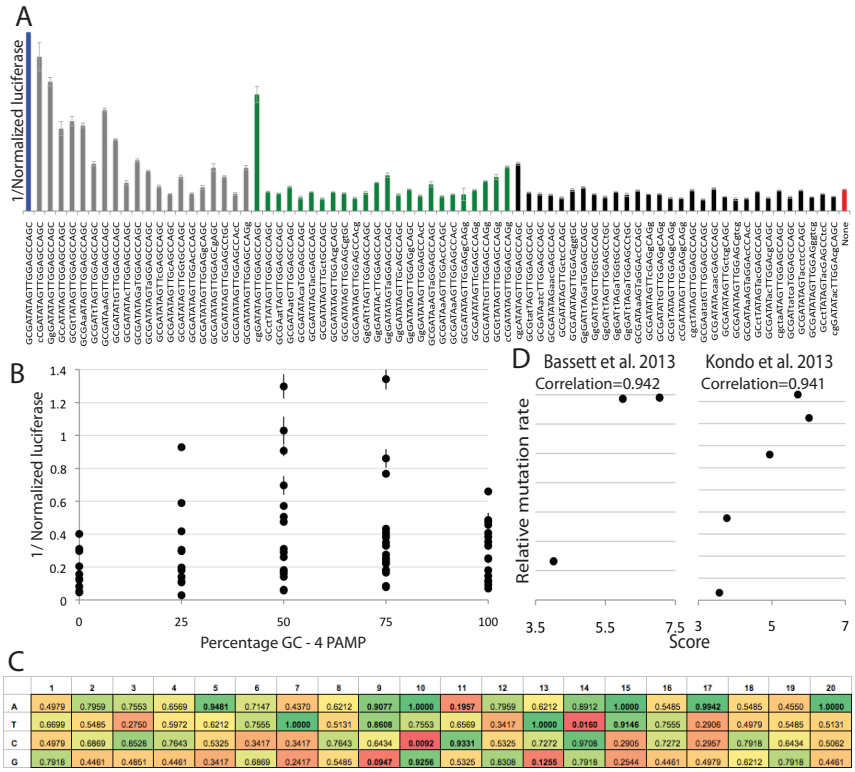
We have made excellent progress at characterizing the phenotype of tumor suppressor in gut stem cells. Specifically, we have shown that single knockdowns of *PTEN*, *AMPK*, *TSC1*, or *TSC2* in *Drosophila* adult gut intestinal stem stems (ISCs), via expression of shRNAs from an ISC-specific promoter (esg), lead to the rapid generation of hyperproliferative lesions/tumors (**Figure 5A**). In addition, we have developed a blood cell assay that will allow us to examine the effect of shRNA combinations on blood cell proliferation (**Figure 5B).** These assays will be used to validate the synthetic lethal pairs that have emerged from Task 1.

Task 3: Determine whether synthetic lethal combinations found in *Drosophila* are relevant to mammalian networks. (months 1-36)

- Characterize the signaling and growth properties of the knockout MEFs for all of the five major tumor suppressor genes of interest (LKB1, NF1, PTEN, TSC1, and TSC2). (months 1-8) COMPLETED
- Identify the mammalian orthologs using DIOPT. (months 8-30) UNDERWAY
- Characterize effective siRNAs against the genes to be targeted using ON-TARGETplus SMARTpool siRNAs from Dharmacon. (months 12-30) UNDERWAY
- Test the synthetic lethal interactions with the RTK network core tumor suppressors in MEFs. (months 12-30) UNDERWAY
- Confirm the results with neutral base pair substitutions control siRNAs. (months 12-30)
- Test hits that show specificity for killing of one or more tumor suppressor-deficient cell types using available tumor-derived cell lines lacking these tumor suppressors. (months 26-36)
- Test all synthetic lethal interactions, regardless of the MEF results, in a TSC2 null angiomyolipoma-derived cell line. (months 26-36)

With the *Drosophila* TSC1 and TSC2 knockout lines now generated and the initial synthetic lethal screens completed, we are testing the hits in mammalian cell lines. We have used the DIOPT software to identify the human orthologs of the high confidence hits from the synthetic lethal screen. Our plan is to test all such hits, as they are verified in the fly system. We are currently testing siRNAs against those targets that scored as selectively kill both TSC1 and TSC2 null cells, as an added stringency filter for specific targets. The identified orthologs include: 1) RNGTT (mRNA-CAP), which is the RNA guanyltransferase and 5'-phosphatase that initiates the capping of mRNAs; 2) CDK11B (Pitslre), a member of the cyclin-dependent kinase family implicated in the control of both mitosis and autophagy; 3) CCNT1 (CycT), a cyclin family member that, through its association with CDK9, forms the P-TEFb complex involved in activating RNA pol II-mediated transcription. For these studies, we are comparing the effects of knocking down these genes on $Tsc2^{+/+}$ and $Tsc2^{-/-}$ MEFs. This system allows us to test the selectivity for any effects we see on viability, thereby providing an indication of a potential therapeutic window or not. We have validated siRNAs for these targets using qRT-PCR (data not shown) and immunoblotting (**Figure 6A**) to assess the level of knockdown. In the past month, we have generated viability curves using these siRNAs. Thus far, we have found that

both RNGTT and CCNT1 (Cyclin T1) knockdowns show more pronounced effects on the viability of *Tsc2-/-* cells than their wild-type counterparts (**Figure 6B**), thereby providing cross-phyla confirmation of these genetic relationships. Of these, RNGTT seems to be the more promising target. We are now in the process of testing the third *Drosophila* hit (CDK11), for which we have recently identified effective siRNAs against (**Figure 6A**). We will then assess specific effects on cell cycle progression and apoptosis with these three targets. It is anticipated that these studies will be completed in the next two months, at which time we will perform rescue experiments to confirm the specificity of any selective effects on viability that we detect and test these interactions in human cell lines.



C

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0.4979	0.7959	0.7553	0.6569	**0.9481**	0.7147	0.4370	0.6212	**0.9077**	**1.0000**	**0.1957**	0.7959	0.6212	0.8912	**1.0000**	0.5485	**0.9942**	0.5485	0.4550	**1.0000**
T	0.6699	0.5485	0.2750	0.5972	0.6212	0.7555	**1.0000**	0.5131	**0.8608**	0.7553	0.6569	0.3417	**1.0000**	**0.0160**	**0.9146**	0.7555	0.2906	0.4979	0.5485	0.5131
C	0.4979	0.6869	0.8528	0.7643	0.5325	0.3417	0.3417	0.7643	0.6434	**0.0092**	**0.9331**	0.5325	0.7272	0.9708	0.2905	0.7272	0.2957	0.7918	0.6434	0.5062
G	0.7918	0.4461	0.4851	0.4461	0.3417	0.6869	0.2417	0.5485	**0.0947**	**0.9256**	0.5325	0.6308	**0.1255**	0.7918	0.2544	0.4461	0.4979	0.6212	0.7918	0.4461

Figure 1: Optimization of the CRISPR system for use in Drosophila S2R+ cells

A: The graph shows 1/normalized luciferase luminescence to represent relative mutation efficiency for the 75 gRNAs shown on the horizontal axis. The blue bar shows a positive control with a perfect match between the sgRNA and the target sequence. Gray, blue and black bars show sgRNAs with 1, 2 or 3-6 mismatches respectively with varying position within the sequence. The red bar shows a negative control in which no sgRNA was present. B: The scatter plot shows 1/normalized luciferase luminescence (mutation efficiency) on the vertical axis and the percentage GC content in the 4bp proximal to the PAM sequence of the 75 sgRNAs shown. All sgRNAs in this experiment are perfect matches to their respective target sites. C: The results shown in B were used to assess the association of nucleotide content in each position of the 75 sgRNAs with high mutation efficiency. These data were then used to produce the probability matrix shown. Numbers represent the p-value associated with each base in each position leading to high mutation efficiency defined by >50% of maximum efficiency shown in B. D: The matrix shown in C was sued to generate a sequence-based scoring algorithm to predict sgRNA efficiency. To validate this, scores were generated for sgRNA sequences with published mutation efficiencies from two independent studies. In both cases, a strong correlation is observed between score and reported efficiency. Note that sgRNAs from these studies that were unlikely to have representative mutation efficiencies due to position within the target gene or reduced viability were not included in this analysis.

A

Media		Clones	% survival
Schneider's	10% FBS	0	0.0
	15% FBS	0	0.0
	20% FBS	0	0.0
Conditioned Schneider's	10% FBS	30	15.8
	15% FBS	31	16.3
	20% FBS	32	16.8

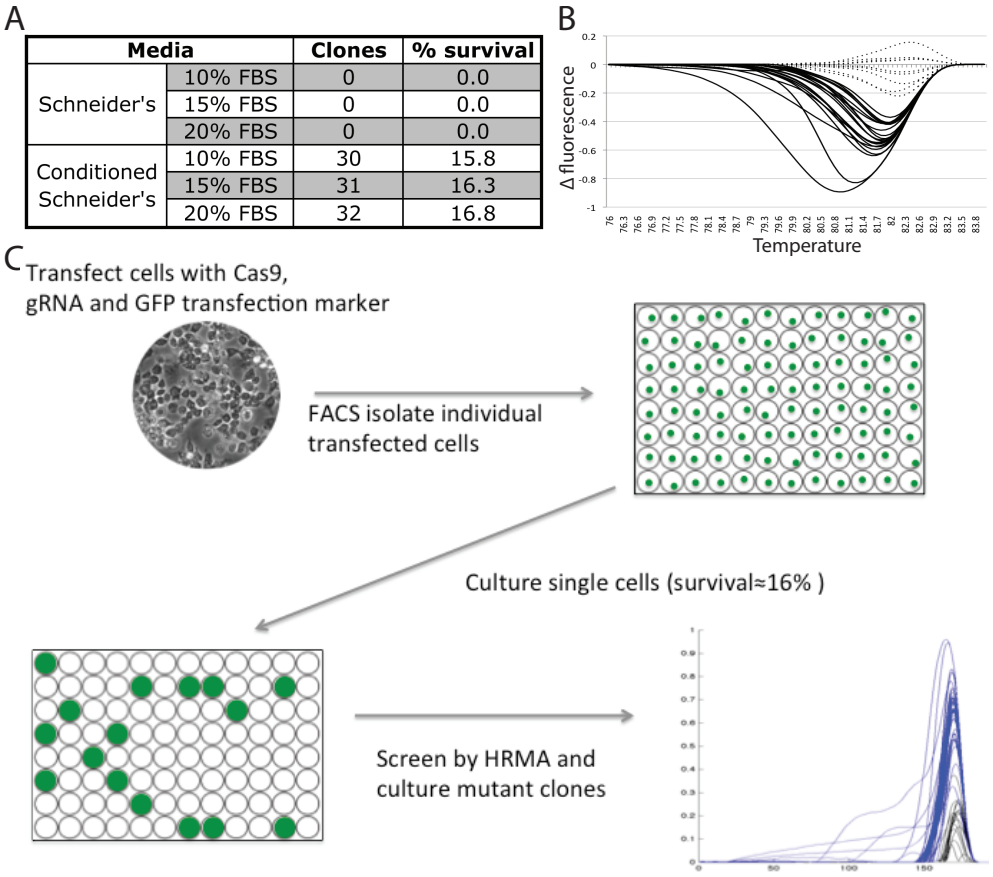


Figure 2: A method to generate stable, homogeneous mutant S2R+ cell lines
A: Several culture media variations were tested to find conditions allowing single cell cloning of S2R+ cells. The table shows the survival rate of single cell clones under 6 different conditions. Conditioned media was produced by culturing wildtype S2R+ cells in fresh media for 16 hours. Cells were then removed by filtration to isolate conditioned media. B: HRMA assays were used to detect single cells from a CRISPR-treated population that carried mutations of the target site. Black lines represent cells differing significantly from controls and dashed lines represent samples that do not differ significantly from controls and are therefore unlikely to carry mutations. C: Schematic of the approach to combine CRISPR-based genome editing and single cell cloning to generate mutant cell lines.

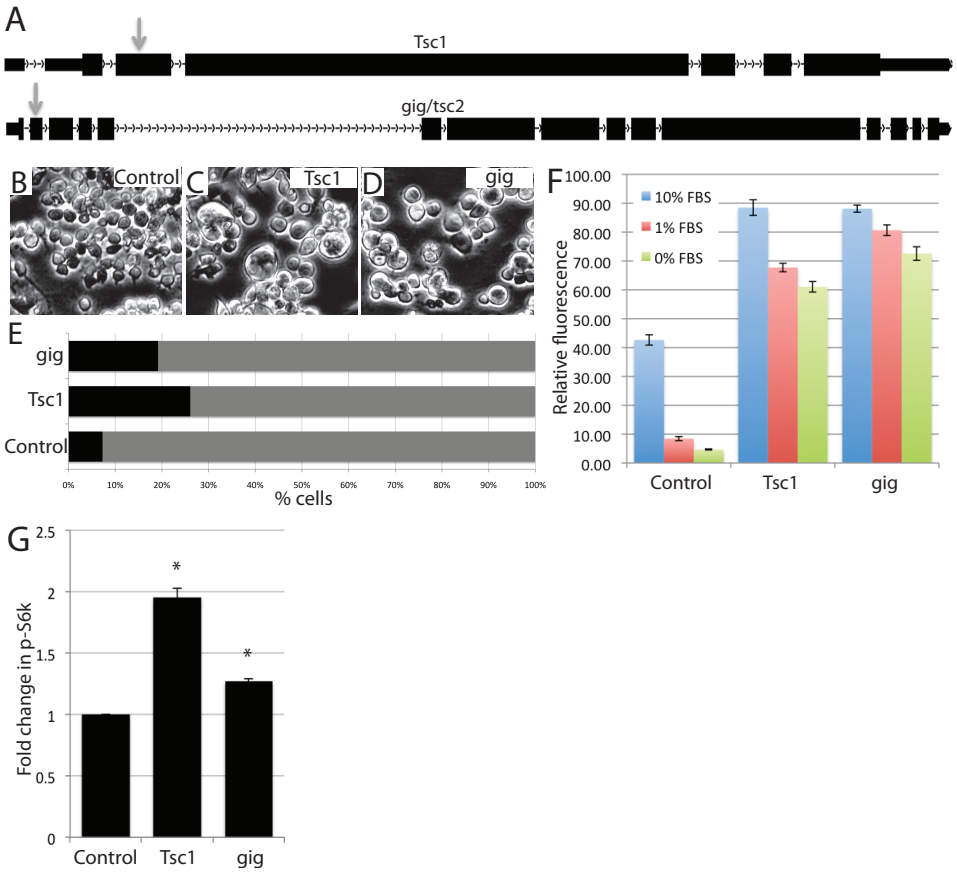


Figure 3: Characterization of Tsc1 and gig/Tsc2 mutant cell lines

A: Schematic indicating the site of mutation in the Tsc1 and gig/Tsc2 genes. B-D: Images of wild type, Tsc1 mutant and gig/Tsc2 mutant cells illustrating an increase in cell size as expected from known phenotypes for these genes. E: Quantification of cell size distribution for each cell line. Black bars indicate the proportion of cells with diameter over an arbitrary threshold. Gray bars indicate the proportion of cells under the same diameter threshold. F: The population growth rates of each cell type were measured after 4 days of culture using CellTitre Glo assays under normal culture conditions (blue bars) or under starvation conditions (red bars – partial starvation and green bars – complete starvation). Results show that the mutant cell lines are no longer responsive to changes in the nutrient status of the culture media.

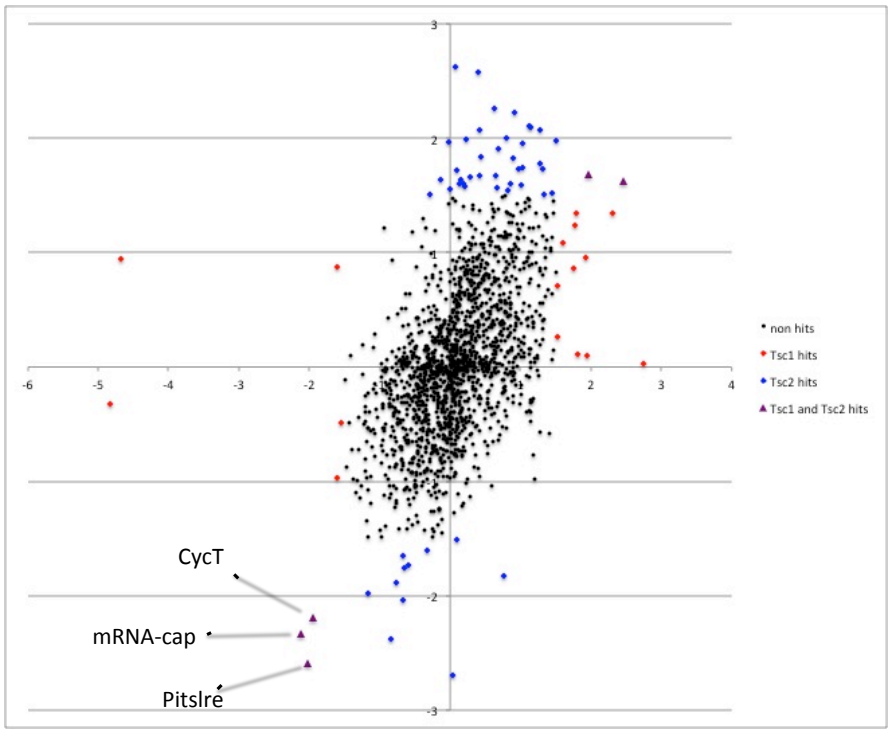


Figure 4: Synthetic screening results
The scatter plot displays Z-scores from the screens in Tsc1 and gig/Tsc2 mutant cell lines. Samples that displayed significant changes in viability in wild type cells are not shown. The three genes that have synthetic lethal relationships with both Tsc1 and gig/Tsc2 are indicated.

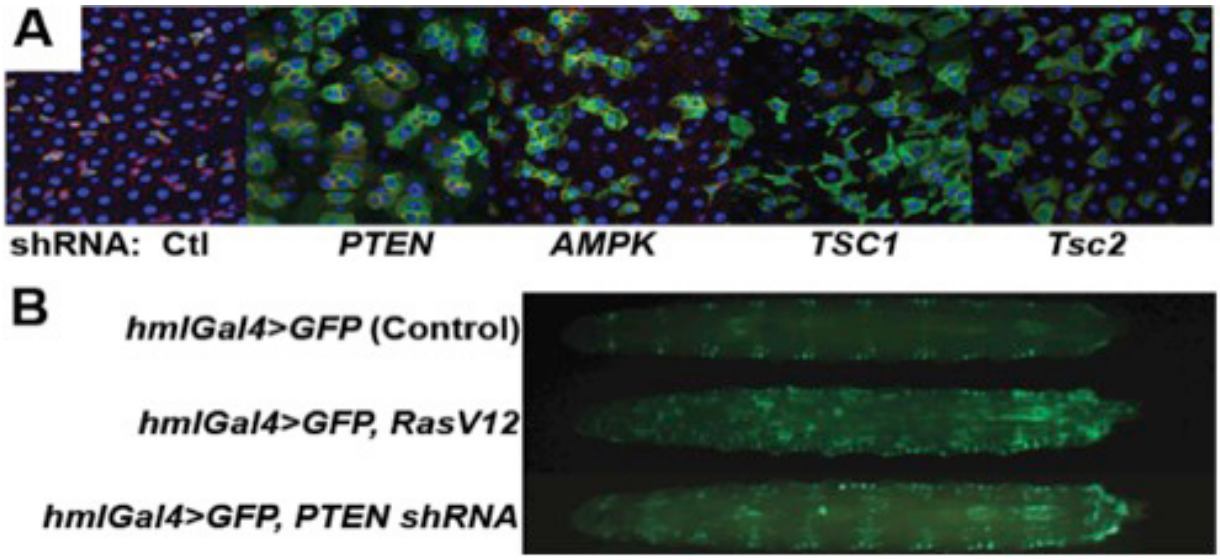


Figure 5: In vivo assays
A: As in mammals, fly ISCs divide about once a day to maintain themselves and produce daughter cells (marked by esg-GFP in green) that differentiate into enterocytes (large blue nuclei). Upon shRNA-knockdown of *PTEN*, *AMPK*, or *TSC1/2* in ISCs (esg-Gal4;UAS-geneX-RNAi), ISCs proliferate and develop "gut tumors" within three days, detectable by GFP. B: Coexpression of RasV12 or PTEN shRNAs with GFP using the hemolectin (hml)-Gal4 driver demonstrates increase larval blood cell proliferation.

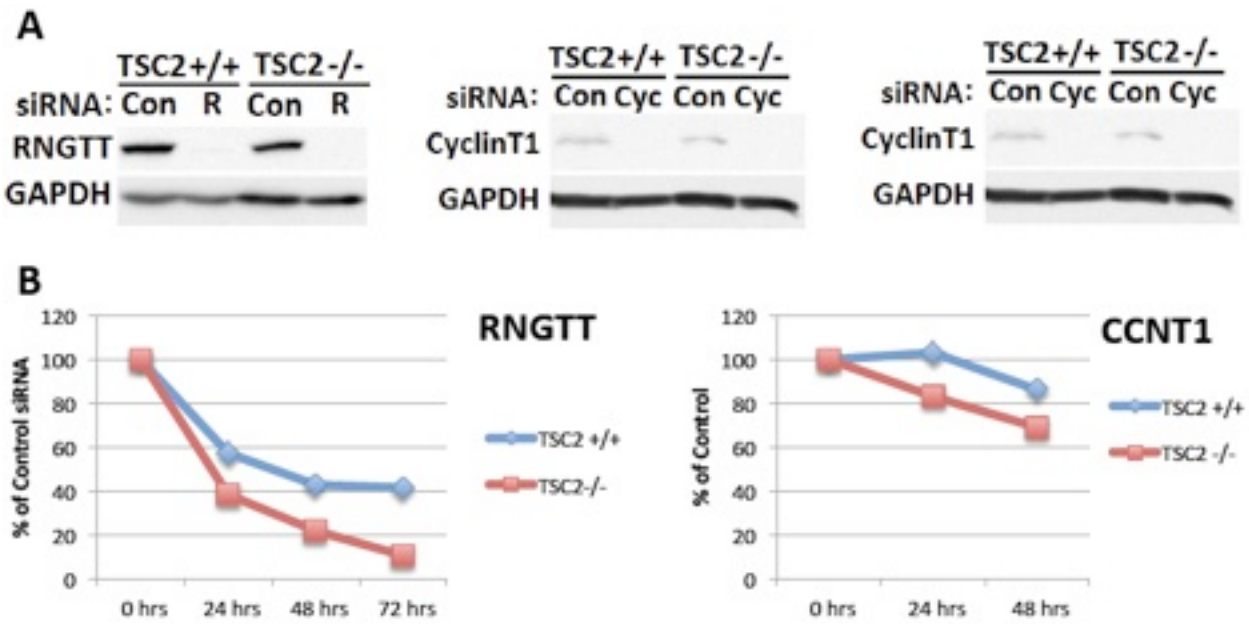


Figure 6: Validation of hits from *Drosophila* synthetic lethal screen in mouse embryo fibroblasts (MEFs). (A) The most effective siRNAs against the three targets identified in the *Dropsophila* genetic screen are shown, demonstrating strong knockdown of the mammalian orthologs by immunoblot with available antibodies. (B) The effects of these siRNAs on the viability of *Tsc2+/+* versus *Tsc2-/-* MEFs are shown, measured by Cell-Titer Glo and normalized to each cell line transfected with control, non-targetting siRNAs.

4. KEY RESEARCH ACCOMPLISHMENTS:

. Successfully produced mutant cell lines for three of the tumor suppressor genes (Nf1, Tsc1 and gig/Tsc2)

. Successfully performed combinatorial screens by treating the mutant lines with dsRNAs targeting all kinases and phosphatases (563 genes) in the *Drosophila* genome.

. Identified three hits (mRNA-Cap, Pitslre and CycT) that scored as synthetic lethal with both Tsc1 and gig/Tsc2 mutations.

. Validated the mammalian ortholog of mRNA-Cap (RNGTT) as a synthetic lethal hit in mammalian cells, selectively decreasing the viability of TSC2 null cells.

5. CONCLUSION:

The tuberous sclerosis complex (TSC) tumor suppressors, Tsc1 and Tsc2, function together in an evolutionarily conserved protein complex that is a point of convergence for major cell signaling pathways regulating mTOR complex 1 (mTORC1). Mutation or aberrant inhibition of TSC is common in a diverse array of human tumor syndromes and cancers across tissue lineages. The discovery of novel therapeutic strategies to selectively kill cells with functional loss of this complex is therefore of significant clinical relevance. We have developed a CRISPR-based method allowing the generation of homogeneous mutant *Drosophila* cell lines. By combining *Tsc1* and *Tsc2* mutant cell lines with RNAi screens against all kinases and phosphatases, we identified synthetic lethal interactions with *Tsc1* and *Tsc2*. Knockdown of two hits in particular (mRNA-cap/RNGTT and CycT/CCNT1) reduced viability of both *Drosophila Tsc1* and *Tsc2* mutant cells but left wild type cells unaffected. Importantly, knockdown of both genes displayed similar selective viability effects in mammalian TSC2-deficient cell lines, including tumor-derived cell lines from a TSC patient, illustrating the power of this cross species screening strategy to identify new drug targets. The long-term impact of these studies is to identify drug targets for treating patients with TSC tumors.

6. PUBLICATIONS, ABSTRACTS, AND PRESENTATIONS:

Benjamin E. Housden, Shuailiang Lin, Yanhui Hu, Ian Flockhart, Charles Roesel, Colleen Kelley, Michael Buckner, Rong Tao, Bahar Yilmazel, Stephanie Mohr and Norbert Perrimon. Generation of clonal *Drosophila* mutant cell lines.

7. INVENTIONS, PATENTS AND LICENSES:

Nothing to report.

8. REPORTABLE OUTCOMES:

Nothing to report.

9. OTHER ACHIEVEMENTS:

Nothing to report.

10. REFERENCES:

Barry ER, Morikawa T, Butler BL, Shrestha K, de la Rosa R, Yan KS, et al. Restriction of intestinal stem cell expansion and the regenerative response by YAP. Nature. 2013;493(7430):106-10.

Briggs AW, Rios X, Chari R, Yang L, Zhang F, Mali P, et al. Iterative capped assembly: rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. Nucleic Acids Res. 2012;40(15):e117.

Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, et al. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res. 2011;39(12):e82.

Charpentier E, Doudna JA. Biotechnology: Rewriting a genome. Nature. 2013;495(7439):50-1.

Cho SW, Kim S, Kim JM, Kim JS. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat Biotechnol. 2013;31(3):230-2.

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013;339(6121):819-23.

Dahlem TJ, Hoshijima K, Jurynec MJ, Gunther D, Starker CG, Locke AS, et al. Simple methods for generating and detecting locus-specific mutations induced with TALENs in the zebrafish genome. PLoS Genet. 2012;8(8):e1002861.

Gratz SJ, Cummings AM, Nguyen JN, Hamm DC, Donohue LK, Harrison MM, et al. Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease. Genetics. 2013.

Housden, B. E., Lin, S., Hu, Y., Flockhart, I., Roesel, C., Kelley, C., Buckner, M., Tao, R., Yilmazel, B., Mohr, S., and Perrimon, N. (2014). Generation of clonal *Drosophila* mutant cell lines. Submitted.

Liu J, Li C, Yu Z, Huang P, Wu H, Wei C, et al. Efficient and specific modifications of the Drosophila genome by means of an easy TALEN strategy. J Genet Genomics. 2012;39(5):209-15.

Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotechnol. 2013;31(3):227-9.

Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, et al. Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol. 2011;29(8):731-4.

Karpowicz P, Perez J, Perrimon N. The Hippo tumor suppressor pathway regulates intestinal stem cell regeneration. Development. 2010;137(24):4135-45.

Lin G, Xu N, Xi R. Paracrine Wingless signalling controls self-renewal of Drosophila intestinal stem cells. Nature. 2008;455(7216):1119-23.

Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. Science. 2013;339(6121):823-6.

Markstein M, Pitsouli C, Villalta C, Celniker SE, Perrimon N. Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. Nat Genet. 2008;40(4):476-83.

Micchelli CA, Perrimon N. Evidence that stem cells reside in the adult Drosophila midgut

epithelium. Nature. 2006;439(7075):475-9.

Moore FE, Reyon D, Sander JD, Martinez SA, Blackburn JS, Khayter C, et al. Improved somatic mutagenesis in zebrafish using transcription activator-like effector nucleases (TALENs). PLoS One. 2012;7(5):e37877.

Ohlstein B, Spradling A. The adult Drosophila posterior midgut is maintained by pluripotent stem cells. Nature. 2006;439(7075):470-4.

Ohlstein B, Spradling A. Multipotent Drosophila intestinal stem cells specify daughter cell fates by differential notch signaling. Science. 2007;315(5814):988-92.

Radtke F, Clevers H. Self-renewal and cancer of the gut: two sides of a coin. Science. 2005;307(5717):1904-9.

Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK. FLASH assembly of TALENs for high-throughput genome editing. Nat Biotechnol. 2012;30(5):460-5.

Sanjana NE, Cong L, Zhou Y, Cunniff MM, Feng G, Zhang F. A transcription activator-like effector toolbox for genome engineering. Nat Protoc. 2012;7(1):171-92.

Zu Y, Tong X, Wang Z, Liu D, Pan R, Li Z, et al. TALEN-mediated precise genome modification by homologous recombination in zebrafish. Nat Methods. 2013;10(4):329-31.

11. APPENDICES:

See next.

Generation of clonal *Drosophila* mutant cell lines

Benjamin E. Housden[1], Shuailiang Lin[1], Yanhui Hu[1,2], Ian Flockhart[1,2], Charles Roesel[1,2,3], Colleen Kelley[1], Michael Buckner[1,2], Rong Tao[1], Bahar Yilmazel[1,2], Stephanie Mohr[1,2] and Norbert Perrimon[1,2,4]

[1] Department of Genetics

[2] *Drosophila* RNAi Screening Center

[3] Northeastern University

[4] Howard Hughes Medical Institute

Harvard Medical School

Boston, MA 02115, USA

ABSTRACT

Drosophila is a well-established *in vivo* system, yet for many experiments such as high-throughput screens and biochemical assays, cell lines of specific genotypes would be invaluable. CRISPR has recently been shown to be an effective tool for genome engineering of mammalian cell lines, however applying this approach to *Drosophila* cells is challenging as they are aneuploid and not easily clonable. Here, we overcome these limitations by developing a method for single-cell cloning and an online tool for quantitative detection of mutations in complex cell culture samples based on high-resolution melt data. We applied these tools to generate and characterize mutant cell lines for *STAT92E*, *Tsc1* and *gig/tsc2* for cell signaling studies. In addition, we established *Lig4* mutant cell lines and show that they enable high-efficiency engineering by homologous recombination. Finally, we performed an in depth analysis of the specificity and efficiency of gRNAs and establish rules for the simple selection of high efficiency gRNAs with no predicted off-targets.

INTRODUCTION

In contrast to the sophisticated toolkit available for genetic studies *in vivo*[1], methods available to manipulate cell lines in *Drosophila* are relatively limited[2]. Current methods to perturb gene functions are based on transfection of various constructs, establishment of stable cell lines, and delivery of RNAi reagents.

Recent advances in genome engineering technologies such as CRISPR have revolutionized our ability to modify genomic sequences. Such systems can be used to generate a double strand break at a defined genomic locus. This is then repaired either by the non-homologous end-joining (NHEJ) pathway, generating small insertion or deletion (indel) mutations, or by homologous recombination (HR) with a donor construct, resulting in a defined sequence change[3-6]. Studies in mammalian cell culture have demonstrated that CRISPR can be used to induce mutations with high efficiency[7-9]. Thus, given the success of CRISPR to induce homozygous mutations efficiently in mammalian cells and its apparently widespread functionality, we explored whether CRISPR would enable the efficient engineering of *Drosophila* cell lines. In particular, mutant cell lines associated with complete loss of gene activity are of special interest, as they would facilitate biochemical studies, in particular of signal transduction mechanisms, as well as combinatorial loss of function screens to identify genes that act redundantly or synergistically. Combinatorial screens are currently performed by cotransfecting cells with multiple RNAi reagents[10, 11]. However, both performing and analysing the results of 'double-knockdown' RNAi screens can be challenging, *e.g.* due to incomplete transfection efficiency and off-target effects associated with RNAi reagents. An alternative approach would be to perform single RNAi screens in mutant cell lines.

However, the generation of mutant *Drosophila* cell lines is particularly challenging as the majority of established cell lines are aneuploid[12], making the efficient generation of homozygous mutant cells in a population considerably more difficult than in diploid mammalian cells. In addition, no methods currently exist for

robust single-cell cloning of *Drosophila* cells, which is required to isolate a clonal culture carrying the desired sequence.

To apply CRISPR technology to the generation of *Drosophila* mutant cell lines, we developed a method for detection of mutations in complex cell culture samples and a method for cloning single *Drosophila* cells. Specifically, we developed HRMAnalyzer, an online tool for quantitative detection of mutations in complex cell culture samples based on high-resolution melt (HRM) data. In addition, we developed an optimized approach to single-cell cloning, greatly improving efficiency relative to previous techniques and for the first time allowing the robust generation of clonal cell populations. Here, we have applied these tools to generate *STAT92E*, *Tsc1* and *gig/tsc2* mutant cell lines that can be used for signal transduction studies. Further, we show that genome modification by HR can be performed with highly improved efficiency in a *Lig4* mutant cell line, in which the NHEJ pathway is inhibited, therefore favoring DNA repair by HR[4]. Finally, in order to optimize the above methods, we present in depth analysis of CRISPR specificity by testing the mutation rates of many different gRNAs against a single target with varying numbers and positions of mismatches, and present a sequence based scoring scheme allowing the accurate prediction of gRNA efficiency.

RESULTS

HRMAnalyzer: An online tool for mutation detection and analysis

Recent advances in genome editing technologies such as TALENs and CRISPR have produced a growing need for efficient and cost-effective methods to detect and analyze changes in genomic sequences. Whilst several approaches exist to detect induced mutations, including nuclease assays[7, 13], sequencing and high-resolution melt (HRM) assays[14, 15], each is associated with limitations. HRM assays offer the potential for cost-efficient and high-throughput screening for mutations not matched by other available methods. The main limitations of HRM are the

requirement for specialized commercial software for data analysis and the inability to perform quantitative analysis of mutation rates. To address these issues, we developed an online tool (HRMAnalyzer; www.flyrnai.org/hrma) for analysis of HRM data generated on standard qPCR machines.

With this approach, normalization steps are first performed to correct for differences in DNA concentrations and noise between samples (**Fig. 1A**). Existing HRM data processing software generally uses clustering-based analysis to divide samples into groups likely to carry similar sequences and therefore identify mutations relative to the control group[16] (see user guides for details on specific commercial software approaches). However, this method is best suited for detection of mutations in relatively simple samples, such as those obtained from single diploid animals, and performs less well on complex mixtures of mutations[17] (compare Fig. 1D to 1B). Instead, HRMAnalyzer uses statistical analysis to identify curves that differ significantly from the distribution of control samples (**Fig. 1A**), allowing the detection of mutations even in highly complex samples.

To test the ability of HRMAnalyzer to detect simple differences in sequence in *in vivo* samples, we performed HRM assays on a fragment amplified from the hop^{TumL} allele of the *hopscotch* (*hop*) gene, which carries a single base pair substitution compared to wild type. The assay was performed in parallel on both wild type flies and flies heterozygous for hop^{TumL} and, using the wild type samples as controls, statistical analysis resulted in significant p-values for all samples derived from hop^{TumL} flies (**Fig. 1B**).

Next, to assess the performance of HRMAnalyzer on more challenging samples, we tested its ability to detect mutations when present at lower frequencies but still with only two different sequences present in the sample. The y^1 mutation of the *yellow* gene consists of a single base pair insertion within the start codon, resulting in yellow-bodied flies. Genomic DNA prepared from y^1 flies was serially diluted using wild type genomic DNA such that samples were produced with two fold

dilutions up to 1024 fold (i.e. 1 mutant copy of the locus per 2048 molecules). HRM was performed in parallel on each of these dilutions and analyzed using HRMAnalyzer. Although differences in melt curves could no longer be easily distinguished visually (**Fig. S1A**), all samples could be successfully differentiated from control curves using HRMAnalyzer (**Fig. 1C**). Similar results were obtained using serially diluted *hop*^{TumL} samples (**Fig. S1B**).

Finally, we tested whether mutations could be detected in complex samples containing many different sequence alterations with variable frequencies, such as in samples obtained from cultured cells. To achieve this, we developed a vector in which a *Drosophila* codon-optimized Cas9 gene is expressed under the control of the *actin5c* promoter and a gRNA is expressed from the *Drosophila* U6b promoter (pl018) (**Supplementary file 1**). S2R+ cells were then transfected with pl018 expressing a gRNA targeting the *yellow* gene. HRMAnalyzer was used to analyze genomic DNA extracted from populations of treated cells containing many different mutations. Similar to the *in vivo* assays, significant p-values were obtained for all CRISPR-treated samples, consistent with detection of mutations in these highly complex samples (**Fig. 1D**).

Quantitative analysis of HRM data

In order to generate p-values associated with each experimental curve, HRMAnalyzer measures the overall difference between samples by calculating the area between each curve and the mean control curve (total area). In simple samples, this value will depend on several factors, including the size and type of mutations present, relative proportion of heterodimers and homodimers and assay noise. However, we hypothesized that in complex samples, due to the large number of different mutations present in the sample, total area would be based only on the proportion of mutated alleles and noise. Assuming that the former is considerably

larger than the latter, this value could be used to extract relative mutation rates between samples.

To test this, we first generated a quantitative mutation reporter system that allows us to compare HRM-derived data with a functional readout. A gRNA target from the *yellow* gene was cloned immediately downstream of the start codon of the luciferase gene driven by an inducible promoter (metallothionein) such that frameshift mutations at the target site would reduce luciferase levels (**Supplementary file 2**). This reporter was then co-transfected into S2R+ cells with 75 different gRNAs containing various levels of mismatch to the target site in order to vary mutation efficiency. Samples were incubated for 24 hours to allow accumulation of mutations before induction of reporter expression and measurement of luciferase activity. For comparison, we then used HRMAnalyzer to detect mutations of the endogenous copy of the target site after a similar period of mutagenesis, allowing us to compare results obtained using the functional readout to the quantitative analysis tool. Values for total area were generated and plotted against mutation rate data from the luciferase reporter system (**Fig. 1E**). Overall, results from these two assays correlate well (correlation coefficient=0.91), indicating that HRM can be used for quantitative assessment of mutation rates in complex samples.

Generation of mutant *Drosophila* cell lines

Generation of a mutant cell line requires not only efficient and specific induction of mutations at the desired locus but also a method to homogenize the culture. Otherwise wild type or ineffective mutations (e.g. for knockout approaches, non-frameshift mutations) will likely remain in the population. The presence of such mutations would lead firstly to weaker phenotypes than a homogenous mutant population and secondly, when the desired mutation causes a proliferative or survival disadvantage, a rapid loss of the mutation from the population[17]. This is a particular problem in *Drosophila* S2R+ cells because the high level of aneuploidy means that

each locus is present between four and six times[12]. Thus, a relatively low frequency of cells will carry suitable mutations on all alleles.

One solution to this problem is to clone single cells from the population and then screen multiple clones for those that carry only the desired mutations. However, there is currently no robust, established method to generate single-cell clones of *Drosophila* cells and previous attempts using feeder cells, an approach commonly used for mammalian cells, have resulted in very low survival rates[18]. We first tested the dilution approach described in Wheeler et al[19], however this approach did not allow us to reliably recover clones derived from single fly cells. Thus, we decided to take a different approach and ask if we could identify alternative culture conditions supportive of single-cell growth. To do this, we first isolated single cells using fluorescence activated cell sorting (FACS) and then tested their ability to proliferate in various culture media. We individually seeded 190 single S2R+ cells into each of six different formulations of culture media. Half of the samples were seeded into media previously conditioned using S2R+ cells in log phase growth and half into fresh media. Each of these conditions was further divided into three different concentrations of FBS (10%, 15% or 20%). Following three weeks of culture, each clone was analyzed for colony growth. As expected, seeding of cells into standard Schneider's media was insufficient to support single-cell survival and additional supplementation with FBS had no effect. However, seeding of cells into conditioned media resulted in 15.8-16.8% survival, with little difference between FBS concentrations (**Fig. 2A**).

Having identified conditions allowing the robust cloning of single cells, we next tested whether a mutant cell line could be generated by treating cells with CRISPR and then using FACS to isolate cells in which all alleles of the target gene were mutated. We co-transfected S2R+ cells with pl018 plasmid targeting a region of the *yellow* gene and a plasmid driving constitutive expression of GFP to mark transfected cells. After four days, FACS was used to isolate 95 single cells from the

subset of the population with the top 10% of GFP expression. These samples were sorted directly into lysis buffer and genomic DNA extracted from each cell. Fragments surrounding the CRISPR target site were successfully amplified from 30 of the samples and 21 (70%) of these were found to carry mutations using HRM assays (**Fig. 2B, solid lines**). PCR products from 8 of the most significant samples in the HRM assays were cloned and sequenced to determine whether any contained mutations on all alleles (**Fig. 2C**). No wild type sequences were found in any of the 8 sequenced samples. Surprisingly, 6 of them contained a single mutant sequence, suggesting that gene conversion occurs rapidly following induction of mutations, similar to previous observations following TALEN or CRISPR induced mutations in other systems[9, 20-22]. Given the high rate of homozygous mutant cells recovered, the combination of this method with single-cell cloning is likely to represent a robust method for the generation of mutant cell lines (**Fig. 2D**).

Generation and characterization of mutant cell lines

Having developed a method allowing the generation of homogeneous mutant cell populations, we produced two lines with mutations in the *STAT92E* or *Ligase4* (*Lig4*) genes (**Fig. S2**) to test whether protein function could be ablated. To do this, we generated gRNAs targeting a region close to Y711 in *STAT92E* (a tyrosine residue phosphorylated by Hopscotch that is essential for STAT92E transcription factor activity[23]), or close to the 5' end of the coding sequence of *Lig4* (**Fig. 3A**). We then used the approach described above to isolate mutant lines that contained a frameshift deletion including Y711 in *STAT92E* or carried a frameshift mutation at the 5' end of *Lig4*.

To determine whether protein function had been lost in STATΔY711 cells, we co-transfected them with a construct driving expression of the *os* gene, a ligand for the JAK STAT pathway, and a reporter of JAK STAT transcriptional activity (*10X STAT-luciferase*). We found firstly that pathway activity in the absence of ligand

induction was 10.8 fold lower in STATΔY711 cells compared to wild type, indicating that background pathway activity had been ablated (**Fig. 3B, blue bars**). Secondly, whereas expression of *os* in wild type cells resulted in robust induction of the luciferase reporter, this response was lost in STATΔY711 cells (**Fig. 3B, red bars**), indicating that the STAT92E protein was no longer functional. Next, we tested whether the mutant cell line produced a stronger phenotype than that observed following RNAi-based knockdown of *STAT92E* in wild type cells. *os* expression was induced in combination with two different dsRNAs targeting *STAT92E* and although the response to *os* was prevented by dsRNA treatment in wild type cells, reporter activity levels under these conditions were still considerably higher than in untreated STATΔY711 cells (**Fig. 3B, green and purple bars**). This suggests that some STAT92E protein function remained in the presence of dsRNA.

Removal of *Lig4 in vivo* has been shown to increase the efficiency of HR-dependent genome modification by ablating the alternative NHEJ DNA repair pathway and *Lig4* mutant flies have proved to be a useful tool for *in vivo* knock-in experiments[4]. We therefore tested whether a similar knockout in cell lines could be used to increase the rate of HR for knock-in applications in this system. Wild type or *Lig4* cells were co-transfected with plasmids expressing CRISPR components targeting the 5' end of the *expanded* (*ex*) or *Myosin 31DF* (*Myo31DF*) genes and a donor construct containing GFP flanked by 1kb regions homologous to the target site. Donors were designed such that, following HR, GFP coding sequence would be inserted immediately following the start codon of the endogenous gene and therefore should be expressed only if inserted at the correct locus. Three days following transfection, a cell analyzer was used to determine the proportion of cells expressing GFP. For *ex* and *Myo31DF* samples, GFP expression was detected in 20.5% and 11.1% of S2R+ cells respectively, compared to only 2.2% and 2.1% of cells when gRNA was omitted, indicating that GFP expression is due to HR dependent genome

modification. We note that the frequency of HR events that we observe is considerably higher than in a recent study[17], possibly due to differences in vector used. Importantly, when similar HR experiments were performed in a *Lig4* mutant background, 30.6% and 20.5% cells expressed GFP respectively (**Fig. 3C**). This increase in efficiency suggests that this line represents a useful tool for HR-based genome modification experiments in cells.

A further application of mutant cell lines is in the generation of *in vitro* models of specific mutations relevant to human disease, an approach that will facilitate high-throughput and biochemical analyses not possible *in vivo.* The TSC complex is a point of convergence of multiple signaling pathways that is mutated or aberrantly regulated in a high proportion of human cancers[24]. We therefore produced cell lines mutant for two core components of this complex, *Tsc1* and *gig/tsc2* (**Fig. S2**) and analyzed how well they reflected the known *in vivo* phenotypes in flies and mammalian models.

Tsc1 and Gig regulate cell growth and proliferation in response to nutrient levels upstream of the *Tor* complex and mutations have been shown to result in changes in phosphorylation of the Tor target S6k and an increase in cell size *in vivo*[25, 26]. We therefore analyzed *Tsc1* and *gig* mutant lines to determine whether the expected phenotypes were present. First, we analyzed levels of p-S6k and, as expected, they were increased in both *Tsc1* and *gig* mutant cell lines relative to wild type cells (**Fig. 3D**).

Next, we analyzed cell size visually and observed an increase in diameter consistent with results from previous RNAi-based experiments in cells[27] (**Fig. 3E-G**). To compare these differences quantitatively, we used a cell analyzer to divide the cell populations into two groups (low diameter and high diameter) based on forward scatter as a proxy for cell size. We then chose a threshold at which the majority of S2R+ cells fell into the low diameter category. Under these conditions, 28.8% of *Tsc1*

cells and 16.4% of *gig* cells fall into the high diameter category compared to only 7.3% of wild type cells (**Fig. 3H**).

Finally, we measured the growth of the cell populations and their responsiveness to changes in nutrient levels. In the absence of Tsc1 or Gig, we observed a higher rate of population growth, consistent with previously observed *in vivo* phenotypes[28, 29] (**Fig. 3I**). Furthermore, whilst the growth rate of a wild type population decreased by over 75% when nutrient levels were reduced, this change was greatly reduced in *Tsc1* and *gig* populations. This suggests that nutrient-dependent growth regulation is dysfunctional, consistent with disruption of the TSC network.

Optimization of the CRISPR system

Analysis of off-target effects in a cell culture system

One limitation of the CRISPR system in mammals is the prolific occurrence of off-target mutations[30-34]. We therefore sought to assess the extent to which this is likely to be an issue in *Drosophila* cell culture. To analyze this, we compared mutation rates of 75 different gRNAs targeting a single target sequence cloned into the quantitative luciferase reporter vector. These gRNAs carry varying numbers and positions of mismatches to the target sequence and so can be used to assess the relative specificity of the system. Results demonstrate that in some cases, either one or two mismatches are sufficient to prevent mutation of the target site. However, when the mismatches were present at positions close to the 5' end of the gRNA, they were generally well tolerated (**Fig. 4A**, grey and green bars). When three or more mismatches were present, mutations were detected in only one case, when the three mismatches were clustered at the 5' end and even then, mutation rate was only 27% of that observed in the control (**Fig. 4A,** black bars).

Development of a gRNA sequence-based scoring scheme to predict efficiency

Although CRISPR has been shown to be widely functional in many systems, the efficiency with which mutations are induced varies greatly between different gRNA sequences[14, 35, 36]. It has been suggested that the GC content of the gRNA sequence may be associated with efficiency, especially in the 4 nucleotides adjacent to the PAM sequence[36-38]. We therefore took advantage of our newly developed luciferase reporter system to test this correlation. We generated 75 different gRNAs and corresponding reporters with varying GC content and co-transfected these constructs into S2R+ cells. Analysis of mutation rate showed a wide range of different efficiencies. However, no clear correlation between GC content and efficiency was apparent, either considering the whole gRNA sequence or just the four PAM proximal nucleotides (**Fig. 4B and Fig. S3**). However, it appeared that gRNAs with very high or very low GC content consistently displayed low efficiency, consistent with previous observations in human cells[36].

To further investigate factors that may be related to mutation rate, we analyzed the dataset by splitting the gRNA sequences into high and low efficiency groups (**Table S1**) to determine nucleotide enrichment at all 20 positions within the high efficiency group (**Fig. 4C**). By comparing specific gRNA nucleotide compositions to this matrix of enrichments, scores can be generated for any gRNA sequence. To test the applicability of this scoring system outside our own data, we analyzed results from two previous *Drosophila* studies reporting mutation rates associated with specific gRNAs *in vivo*[14, 35]. gRNAs were selected from these studies that had no apparent effect on viability and that were likely to produce a robust phenotype (i.e. those with low numbers of animals recovered or with target sites close to the 3' end of target genes were removed). We assigned scores to each of these gRNAs based on their sequence and compared them to the reported efficiencies (**Fig. 4D**). In both cases, we found a strong correlation (correlation coefficient=0.942 (Bassett) and 0.941 (Kondo)), indicating that the rules identified in our analysis are good predictors

of gRNA function and are likely to be informative in the design of gRNAs for future studies both *in vivo* and in cultured cells.

An updated tool for gRNA design

We previously developed an online tool facilitating the design of gRNAs and prediction of their off-target sites based on rules derived from mammalian studies[39]. However, based on our analysis in *Drosophila* cell lines, these rules appear to be overly stringent. Thus, an improved version of the design tool (CRISPR2) was implemented such that the user can select a stringency cutoff of 3, 4 or 5 mismatches for potential off-target sites (www.flyrnai.org/crispr2). Note that gRNAs annotated in this tool require the seed region (the fifteen positions adjacent to the PAM sequence) to be unique and so annotated gRNA target sites will never have off-targets in which all mismatches are at the 5' end. Therefore, even with the most relaxed stringency settings unannotated off-target sites are unlikely.

In addition, to enable the use of the efficiency scoring system, the online design tool was extended to include efficiency prediction scores for all gRNA targets in the *Drosophila* genome and an additional stand-alone tool to calculate scores for any user defined sequences (www.flyrnai.org/evaluateCrispr). Furthermore, to facilitate selection of a suitable molecular screening method for rapid identification of indels or other mutagenic events, we annotated all gRNAs in the CRISPR2 tool with any relevant restriction enzyme target sites likely to be disrupted by CRISPR treatment

DISCUSSION

To enable biochemical and functional studies in *Drosophila* cells, we developed a method for single-cell cloning and an online tool for quantitative detection of mutations based on high-resolution melt data. We applied these tools to generate mutant cell lines for *STAT92E*, *Tsc1* and *gig/tsc2*, and demonstrate that

they have the expected cellular phenotypes. In addition, we established *Lig4* mutant cell lines that allow high-efficiency engineering by HR, and made newly identified 'rules' for CRISPR design available through our online CRISPR2 and efficiency tools.

One current limitation of the CRISPR system is the occurrence of off-target mutations as demonstrated in mammalian systems[30-34]. Our analysis of CRISPR specificity in *Drosophila* cell lines indicates that mutant lines can easily be produced with no predicted off-target mutation sites even using relatively relaxed criteria. Indeed, we estimate that 97% of protein coding genes can be mutated using gRNAs with no predicted off targets (**Table S3**).

Another limitation has been an inability to predict gRNA efficiency prior to testing and many groups have reported widely varying mutation rates depending on the gRNA sequence[14, 35, 36]. Some correlation between GC content at the 3' end of the gRNA and efficiency has been suggested[36-38] but our analysis suggests that this property plays a minor role in determining mutation rate, as inefficient gRNAs are common at all GC frequencies. Note that an indirect analysis of mutation rates in human cells also found a similar relationship between GC content and efficiency[36]. Instead, we have developed a scoring system based on gRNA sequence that appears to accurately predict mutation rate in both of the independent datasets that we analyzed. The ability to distinguish efficient and inefficient gRNA target sites will simplify experimental design in a wide range of experiments both *in vivo* and in cells.

In summary, we have developed a method allowing, for the first time, the generation of stable homogenous mutant *Drosophila* cell lines. This advance opens up a range of new experimental applications not previously possible. In particular, we envision that mutant cell lines will facilitate combinatorial screens whereby two genes are interrogated for redundant functions or synergistic interactions. In addition, the generation of mutant cell lines will allow biochemical experiments using extracts depleted of specific components. Finally, as the changes that can be engineered using CRISPR are not limited to coding regions, it will be possible to generate cell

lines whereby specific non-coding sequences are modified. This will allow in particular the analysis of transcriptional regulation *in situ*, an application for which there are currently no available methods.

Acknowledgements

We thank David Sabatini for useful discussion on cell cloning and David Doupé for useful discussion on homologous recombination. This work was supported by NIH (5R01DK088718, 5P01CA120964 and R01GM067761) and DOD (W81XWH-12-1-0179). S.E.M. also receives support from the Dana Farber/Harvard Cancer Center, which is supported in part by NCI Cancer Center Support Grant # NIH 5 P30 CA06516. N.P. is a Howard Hughes Medical Institute investigator.

ONLINE METHODS

Generation of CRISPR expression vector

A *Drosophila* codon optimized Cas9 with 3xFlag tag and NLS elements at both 5' and 3' of Cas9 was synthesized by GenScript and the *Drosophila* U6 promoter and act5c promoter were PCR amplified from fly genomic DNA (**Table S4**). These were used to replace the human codon optimized Cas9, human U6 and CGh promoters respectively of the px330[7] plasmid to yield the pl018 plasmid (**Supplementary file 1**).

gRNA homology sequences were cloned into pl018 using pairs of DNA oligonucleotides, which were annealed and ligated into *BbsI* sites according to a previously described protocol[7].

Generation of mutant cell lines

Transfections

Cells were transfected using Effectene Transfection Reagent (Qiagen) according to manufacturers instructions. For generation of mutant cell lines, we used 360ng of pl018 plasmid and 40ng actin-GFP plasmid as a marker of transfected cells. Transfections were performed in 6 well plates and unless stated otherwise, were incubated for four days at 25°C before further processing.

Conditioned media

S2R+ cells were incubated with fresh Schneider's media supplemented with 10% FBS for 16 hours while in log phase growth. Media was then filtered to remove cells and diluted 50% using fresh media supplemented with FBS to obtain the required final FBS concentration.

Single-cell cloning

Cloning of single cells was performed using fluorescence activated cell sorting (FACS) of GFP marked cells. Untransfected cells were used to determine background fluorescence levels before selecting the top 10% of GFP-expressing cells for isolation. Individual cells were sorted into 96 well plates containing culture media. Following two or three weeks of culture, single cells clones were identified visually and isolated into larger cultures.

HRM assays

PCR fragments were prepared from genomic DNA as described for sequencing analysis. Reaction products were then diluted 1:10,000 before an additional round of PCR amplification using Precision Melt Supermix (Bio-Rad) and nested primers to generate a product <120bp in length (95°C 3min, 50 rounds of [95°C 30sec, 60°C 18s, plate read], 95°C 30sec, 25°C 30sec, 10°C 30sec, 55°C 31sec, ramp from 55°C to 95°C and plate read every 0.1°C). Data was analyzed using HRMAnalyzer, available at www.flyrnai.org/HRMA. See Table S4 for primer sequences.

Sequence verification of clones

Genomic DNA was prepared from cultured cells by resuspension in 100uL of lysis buffer (10mM Tris-HCL pH8.2, 1mM EDTA, 25mM NaCl and 200ug/ml proteinase K) and incubation in a thermo cycler for 1 hour at 50°C followed by denaturation at 98°C for 30 minutes.

Target sequences were cloned by PCR using Phusion high-fidelity DNA polymerase (NEB) according to manufacturer's recommendations and supplemented with an additional 2.5 mM $MgCl_2$ (35 cycles: 96°C, 30s; 50°C, 30s; 72°C, 30s). PCR products were gel purified, cloned into the pCR-Blunt II-TOPO vector (Invitrogen) and transformed into Top10 chemically competent cells (Invitrogen). Following

transformation, single colonies were isolated for sequencing. To assess homozygosity of single-cell samples, a minimum of 5 colonies were sequenced per sample. For identification of mutant cell lines a minimum of 20 colonies were analyzed.

Cell characterization assays

Analysis of STAT92E activity

S2R+ and STATΔY711 cell lines were transfected using Effectene Transfection Reagent (Qiagen) according to the manufacturer's instructions to introduce *os* cDNA cloned into pMK33 expression vector, Renilla expression vector (pRL-TK, promega) and 10X-STAT-luc[40] into experimental samples or pMK33, pRL-TK and 10X-STAT-luc into control samples. RNAi samples included an additional 50ng of dsRNA (DRSC ID: DRSC16870 or DRSC37655) from the dsRNA template collection at the *Drosophila* RNAi Screening Center (DRSC). Cells were transfected for 24 hours before addition of $CuSO_4$ at a final concentration of 140μM and incubation for a further 16 hours. Firefly and Renilla luciferase measurements were performed using a SpectraMax Paradigm Multi-Mode Microplate Detection Platform (Molecular Devices).

Homologous recombination experiments

Wild type S2R+ or Lig4 mutant cells were transfected with gRNAs cloned into pl018 (**Supplementary file 1**) targeting the *ex* or *Myo31DF* genes and donor constructs containing GFP coding sequence flanked by 1kb homology arms (**Table S4**). Cells were transfected for 4 days before analysis of GFP expression using a BD Biosciences LSR Fortessa X-20 cell analyzer.

In-cell westerns

Cells were starved in FBS free Schneider's media for 24 hours in 384-well plates before fixing in 4% formaldehyde in PBS for 20 minutes at room temperature. To permeabilize, fixing solution was removed and cells were washed three times with 1X PBS containing 0.1% triton (PBX) for 10 minutes per wash. Triton washing buffer was removed and cells were blocked with PBX + 5% BSA solution (PBT) for one hour. Cells were incubated with primary antibodies (1:200) in PBT with gentle agitation at 4°C overnight. Next, cells were washed with PBT three times for 10 minutes per wash at room temperature before incubation with secondary antibody solutions (1:200) in PBT for two hours. Cells were finally washed three times with PBT for 20 minutes per wash and placed in PBS for imaging and quantification on a Li-cor Aerius system. Primary antibody was p-S6k (T398) (Cell Signaling Technology) and secondary antibody was Alexa Fluor 680 goat anti-rabbit (Invitrogen). p-S6k levels were normalized to tubulin to control for cell number.

Cell size assays

S2R+, *Tsc1* and *gig* mutant cell lines were analysed using a BD Biosciences LSR Fortessa X-20 cell analyzer to measure forward scatter for each cell as a proxy for cell diameter.

Cell line growth assays

5000 cells for each line were seeded into 384 well plates containing 50µl culture media and incubated at 25°C for 5 days. 27µl of CellTiter-Glo reagent (Promega) was added to each well before reading luminescence using a SpectraMax Paradigm Multi-Mode Microplate Detection Platform (Molecular Devices).

Luciferase-based mutation reporter assays

The luciferase reporter vector was constructed by PCR amplifying the metallothionein promoter from pMK33 and luciferase gene from pGL3 (**Table S4**)

and combining these with annealed oligos containing a gRNA target site (**Table S1, Table S2 and Table S4**) and a custom made cloning vector using Golden Gate assembly.

Luciferase assays were performed by transfecting S2R+ cells with the relevant pl018 plasmid, luciferase reporter and pRL-TK (Promega) (to allow normalization of transfection efficiencies between samples) in 96 well plates using Effectene Transfection Reagent (Qiagen) according to the manufacturer's recommendations. 24 hours after transfection, $CuSO_4$ was added to the cell media at a final concentration of 140μM and cells were incubated for a further 16 hours.

Firefly and Renilla luciferase readings were taken using the Dual-Glo Luciferase Assay System (Promega) according to manufacturer's instructions and a SpectraMax Paradigm Multi-Mode Microplate Detection Platform (Molecular Devices).

Online tools

An improved version of CRISPR design tool was implemented re-using some of the modules developed previously[39]. Besides allowing users to choose different off-target thresholds, this version also displays pre-calculated efficiency score and restriction enzyme annotation. The efficiency score was calculated based on a probability matrix computed using the *in vitro* cell line data. It reflects cumulative p-value for high efficiency of each nucleotide from position 1 to 20 with higher values representing higher efficiency (**Fig. S4**). A user interface allowing efficiency score calculation for user-provided sequences was also developed as part of the improved tool, which dynamically calculates predicted efficiency scores for each input sequence from position 1 to 20 or over a user-defined region (**Fig. S4**).

HRMAnalyzer is written as a series of Matlab programs running under control of CGI front-end implemented in Perl and Javascript. The Matlab programs are compiled as stand-alone executable programs and called from within the Perl CGI back-end script. Both tools are hosted on a shared server provided by the Research

IT Group (RITG) at Harvard Medical School.

REFERENCES

1. Venken, K.J. & Bellen, H.J. Genome-wide manipulations of Drosophila melanogaster with transposons, Flp recombinase, and PhiC31 integrase. *Methods in molecular biology* **859**, 203-228 (2012).
2. Cherbas, L., Moss, R. & Cherbas, P. Transformation techniques for Drosophila cell lines. *Methods in cell biology* **44**, 161-179 (1994).
3. Beumer, K., Bhattacharyya, G., Bibikova, M., Trautman, J.K. & Carroll, D. Efficient gene targeting in Drosophila with zinc-finger nucleases. *Genetics* **172**, 2391-2403 (2006).
4. Beumer, K.J. et al. Efficient gene targeting in Drosophila by direct embryo injection with zinc-finger nucleases. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 19821-19826 (2008).
5. Bibikova, M., Beumer, K., Trautman, J.K. & Carroll, D. Enhancing gene targeting with designed zinc finger nucleases. *Science* **300**, 764 (2003).
6. Bibikova, M., Golic, M., Golic, K.G. & Carroll, D. Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. *Genetics* **161**, 1169-1175 (2002).
7. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-823 (2013).
8. Mali, P. et al. RNA-guided human genome engineering via Cas9. *Science* **339**, 823-826 (2013).
9. Wang, H. et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153**, 910-918 (2013).

10. Bakal, C. et al. Phosphorylation networks regulating JNK activity in diverse genetic backgrounds. *Science* **322**, 453-456 (2008).
11. Horn, T. et al. Mapping of signaling networks through synthetic genetic interaction analysis by RNAi. *Nature methods* **8**, 341-346 (2011).
12. Williams, B.R., Bateman, J.R., Novikov, N.D. & Wu, C.T. Disruption of topoisomerase II perturbs pairing in drosophila cell culture. *Genetics* **177**, 31-46 (2007).
13. Kim, H.J., Lee, H.J., Kim, H., Cho, S.W. & Kim, J.S. Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome research* **19**, 1279-1288 (2009).
14. Bassett, A.R., Tibbit, C., Ponting, C.P. & Liu, J.L. Highly efficient targeted mutagenesis of Drosophila with the CRISPR/Cas9 system. *Cell reports* **4**, 220-228 (2013).
15. Dahlem, T.J. et al. Simple methods for generating and detecting locus-specific mutations induced with TALENs in the zebrafish genome. *PLoS genetics* **8**, e1002861 (2012).
16. Dwight, Z.L., Palais, R. & Wittwer, C.T. uAnalyze: web-based high-resolution DNA melting analysis with comparison to thermodynamic predictions. *IEEE/ACM transactions on computational biology and bioinformatics / IEEE, ACM* **9**, 1805-1811 (2012).
17. Bassett, A.R., Tibbit, C., Ponting, C.P. & Liu, J.L. Mutagenesis and homologous recombination in Drosophila cell lines using CRISPR/Cas9. *Biology open* **3**, 42-49 (2014).

18. Neumuller, R.A. et al. Stringent analysis of gene function and protein-protein interactions using fluorescently tagged genes. *Genetics* **190**, 931-940 (2012).

19. Lindquist, R.A. et al. Genome-scale RNAi on living-cell microarrays identifies novel regulators of Drosophila melanogaster TORC1-S6K pathway signaling. *Genome research* **21**, 433-446 (2011).

20. Bedell, V.M. et al. In vivo genome editing using a high-efficiency TALEN system. *Nature* **491**, 114-118 (2012).

21. Gratz, S.J. et al. Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease. *Genetics* **194**, 1029-1035 (2013).

22. Hockemeyer, D. et al. Genetic engineering of human pluripotent cells using TALE nucleases. *Nature biotechnology* **29**, 731-734 (2011).

23. Yan, R., Small, S., Desplan, C., Dearolf, C.R. & Darnell, J.E., Jr. Identification of a Stat gene that functions in Drosophila development. *Cell* **84**, 421-430 (1996).

24. Menon, S. & Manning, B.D. Common corruption of the mTOR signaling network in human tumors. *Oncogene* **27 Suppl 2**, S43-51 (2008).

25. Nobukini, T. & Thomas, G. The mTOR/S6K signalling pathway: the role of the TSC1/2 tumour suppressor complex and the proto-oncogene Rheb. *Novartis Foundation symposium* **262**, 148-154; discussion 154-149, 265-148 (2004).

26. Huang, J. & Manning, B.D. The TSC1-TSC2 complex: a molecular switchboard controlling cell growth. *The Biochemical journal* **412**, 179-190 (2008).

27. Guertin, D.A., Guntur, K.V., Bell, G.W., Thoreen, C.C. & Sabatini, D.M. Functional genomics identifies TOR-regulated genes that control growth and division. *Current biology : CB* **16**, 958-970 (2006).

28. Kwiatkowski, D.J. & Manning, B.D. Tuberous sclerosis: a GAP at the crossroads of multiple signaling pathways. *Human molecular genetics* **14 Spec No. 2**, R251-258 (2005).

29. Tapon, N., Ito, N., Dickson, B.J., Treisman, J.E. & Hariharan, I.K. The Drosophila tuberous sclerosis complex gene homologs restrict cell growth and cell proliferation. *Cell* **105**, 345-355 (2001).

30. Cradick, T.J., Fine, E.J., Antico, C.J. & Bao, G. CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity. *Nucleic acids research* **41**, 9584-9592 (2013).

31. Fu, Y. et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature biotechnology* **31**, 822-826 (2013).

32. Hsu, P.D. et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature biotechnology* **31**, 827-832 (2013).

33. Mali, P. et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature biotechnology* **31**, 833-838 (2013).

34. Pattanayak, V. et al. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nature biotechnology* **31**, 839-843 (2013).

35. Kondo, S. & Ueda, R. Highly improved gene targeting by germline-specific Cas9 expression in Drosophila. *Genetics* **195**, 715-721 (2013).

36. Wang, T., Wei, J.J., Sabatini, D.M. & Lander, E.S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80-84 (2014).

37. Fujii, W., Kawasaki, K., Sugiura, K. & Naito, K. Efficient generation of large-scale genome-modified mice using gRNA and CAS9 endonuclease. *Nucleic acids research* **41**, e187 (2013).

38. Jao, L.E., Wente, S.R. & Chen, W. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 13904-13909 (2013).

39. Ren, X. et al. Optimized gene editing technology for Drosophila melanogaster using germ line-specific Cas9. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 19012-19017 (2013).

40. Bach, E.A. et al. GFP reporters detect the activation of the Drosophila JAK/STAT pathway in vivo. *Gene expression patterns : GEP* **7**, 323-331 (2007).

FIGURE LEGENDS

Figure 1: An online tool for mutation detection and analysis

A: Pipeline of analysis performed by HRMAnalyzer on melt curve data. **B-D:** Results from HRMAnalyzer. Significant samples are represented by solid lines ($p \leq 0.01$) and all other samples by dashed lines. The graph shows normalized difference in fluorescence from mean control (vertical axis) against relative temperature (horizontal axis). Results were extracted from HRMAnalyzer in table format and plotted in excel to improve clarity. Analysis was performed on genomic DNA extracted from hop^{TumL} (blue lines) or wild type flies (black lines) (B), serial dilutions of genomic DNA extracted from y^{1} flies compared to wild type flies (C) or genomic DNA extracted from CRISPR treated S2R+ cells or wild type S2R+ cells (D). **E:** Scatter plot comparing relative mutation rates from 75 gRNAs homologous to a single target sequences measured using two independent assays. The vertical axis shows fold change in mutation rate detected using a luciferase reporter (1/Firefly luciferase activity normalized to Renilla luciferaes activity to control for transfection efficiency) and the horizontal axis shows fold change in mutation rate detected using HRM to assess total area. Error bars represent standard error of the mean from three biological replicates for each assay. Note that the same luciferase assay data are used in Figure 4A.

Figure 2: Generation of homogenous mutant cell lines

A: Table showing survival rates of single S2R+ cells seeded into different media formulations. 'Clones' represents the number of seeded samples that produced viable populations of cells after three weeks. Schneider's media was supplemented with FBS at the concentrations indicated and was preconditioned using S2R+ cells where indicated (see Methods). **B:** HRMAnalyzer results for single S2R+ cells from a population four days after treatment with CRISPR targeting the *yellow* gene. Details are as described for Figure 1. **C:** Sequencing results from 8 of the highest scoring

samples in the HRM assay shown in C. Nucleotides shown in red represent insertions and red dashes represent deletions relative to the wild type sequence as shown in the top row. At least 5 clones were sequenced from each sample to determine whether lines were homozygous mutants. **D:** Workflow showing the major steps required to generate mutant cell lines and estimates of time required for each.

Figure 3: Generation and characterization of mutant cell lines

A: Schematics of the *STAT92E*, *Lig4*, *Tsc1* and *gig* genes. UTRs are represented by thin black boxes, coding exons by thick black boxes and introns by black lines. Arrows superimposed on introns indicate the direction of transcription. CRISPR target sites for each gene are shown the by grey arrows. **B:** Graph showing relative Firefly luciferase activity normalized to Renilla luciferase activity for either wild type or STATΔY711 cell lines in the presence (red bars) or absence (blue bars) of JAK STAT pathway stimulation (*os* expression) and with stimulation in the presence of dsRNA targeting *STAT92E* (green and purple bars). Bars show the mean from two biological replicates and error bars represent standard error of the mean. **C:** Graph showing the percentage of cells expressing GFP following CRISPR induced recombination to insert GFP into the indicated genes. Results show a comparison between wild type S2R+ cells (blue bars) and *Lig4* mutant cells (red bars). **D:** Quantification of p-S6k levels for the cell lines indicated. Bars represent mean fold change in p-S6k levels normalized to Tubulin levels for 4 replicates in each case. Error bars represent standard error of the mean and asterisks indicate significant differences from control ($p \leq 0.01$) based on t-tests. **E-G** Images of representative fields from wild type (E), *Tsc1* mutant (F) or *gig* mutant (G) cell lines. All images were taken at the same magnification and using the same settings. **H:** Graph showing frequency of cell sizes for the cell lines indicated, divided into 'low diameter' (grey bars) or 'high diameter' (black bars) using a cutoff at which the majority of wild type cells fall into the 'low diameter' category. **I:** Graph showing relative rates of population growth for the cell

lines indicated in either complete media (10% FBS – blue bars), under partial starvation conditions (1% FBS – red bars) or complete starvation conditions (no FBS – green bars). Note that these values represent a combination of cell growth and proliferation. Bars show the mean of at least 24 samples and error bars represent standard error of the mean.

Figure 4: Optimization of the CRISPR system

A: Graph showing relative mutation rates from 75 gRNAs used to target a single sequence cloned into a luciferase reporter (note that the same data are used in Figure 1E). Mutation rate is calculated as 1/Firefly luciferase activity normalized to Renilla luciferase activity to control for differential transfection efficiency. Bars show mean relative mutation rates from three biological replicates using gRNAs with 0 mismatches (blue bar), 1 mismatch (grey bars), 2 mismatches (green bars), ≥3 mismatches (black bars) or in the absence of gRNA (red bar). **B:** Graph showing relative mutation rates, measured using luciferase assays, of 75 gRNAs targeting different sequences (vertical axis), compared to GC content of the 4 nucleotides adjacent to the PAM sequence (horizontal axis). Each point shows the mean value from 3 biological replicates and error bars represent standard error of the mean. **C:** Matrix showing enrichment p-values of each nucleotide in each position amongst high efficiency gRNAs from the 75 described in B. **D:** Validation of efficiency scores generated using the matrix shown in C by correlating score (horizontal axis) with efficiency (vertical axis) from two independent publications.

Figure_1: An online tool for mutation detection and analysis

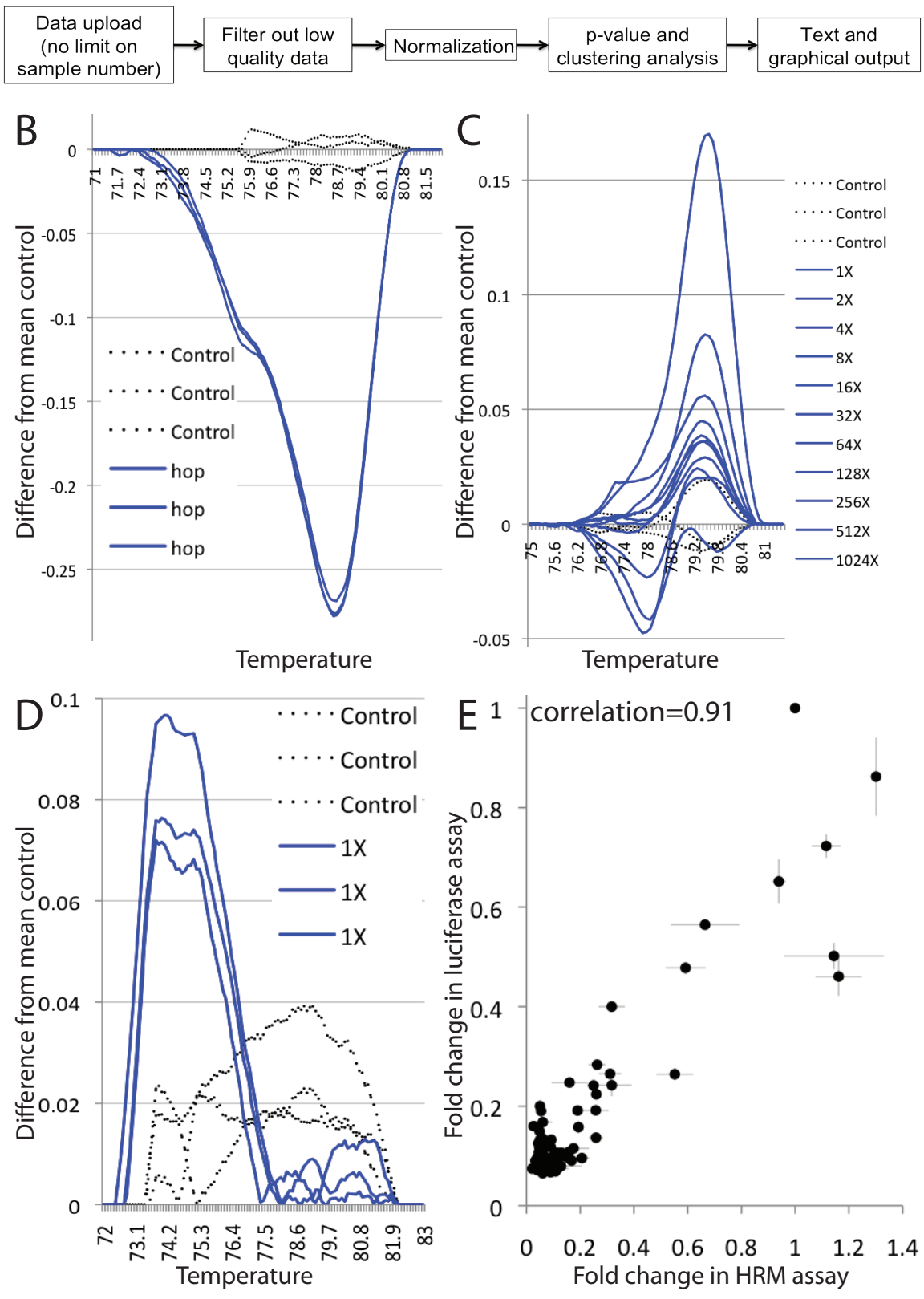
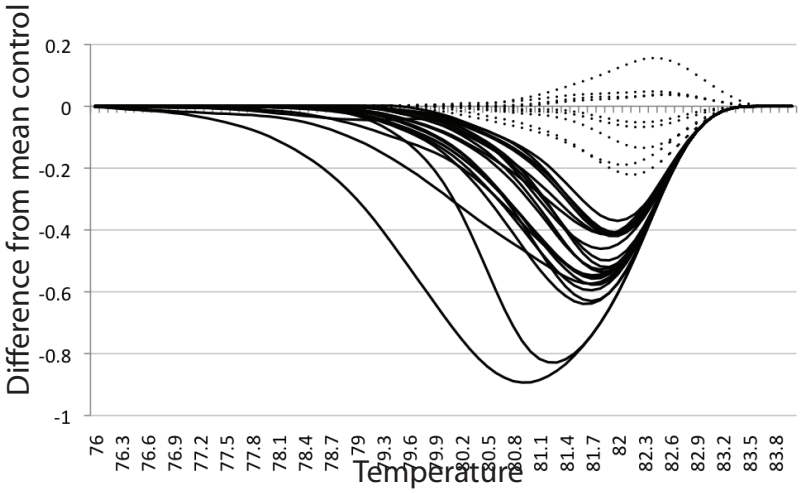


Figure_2: Generation of homogeneous mutant cell lines

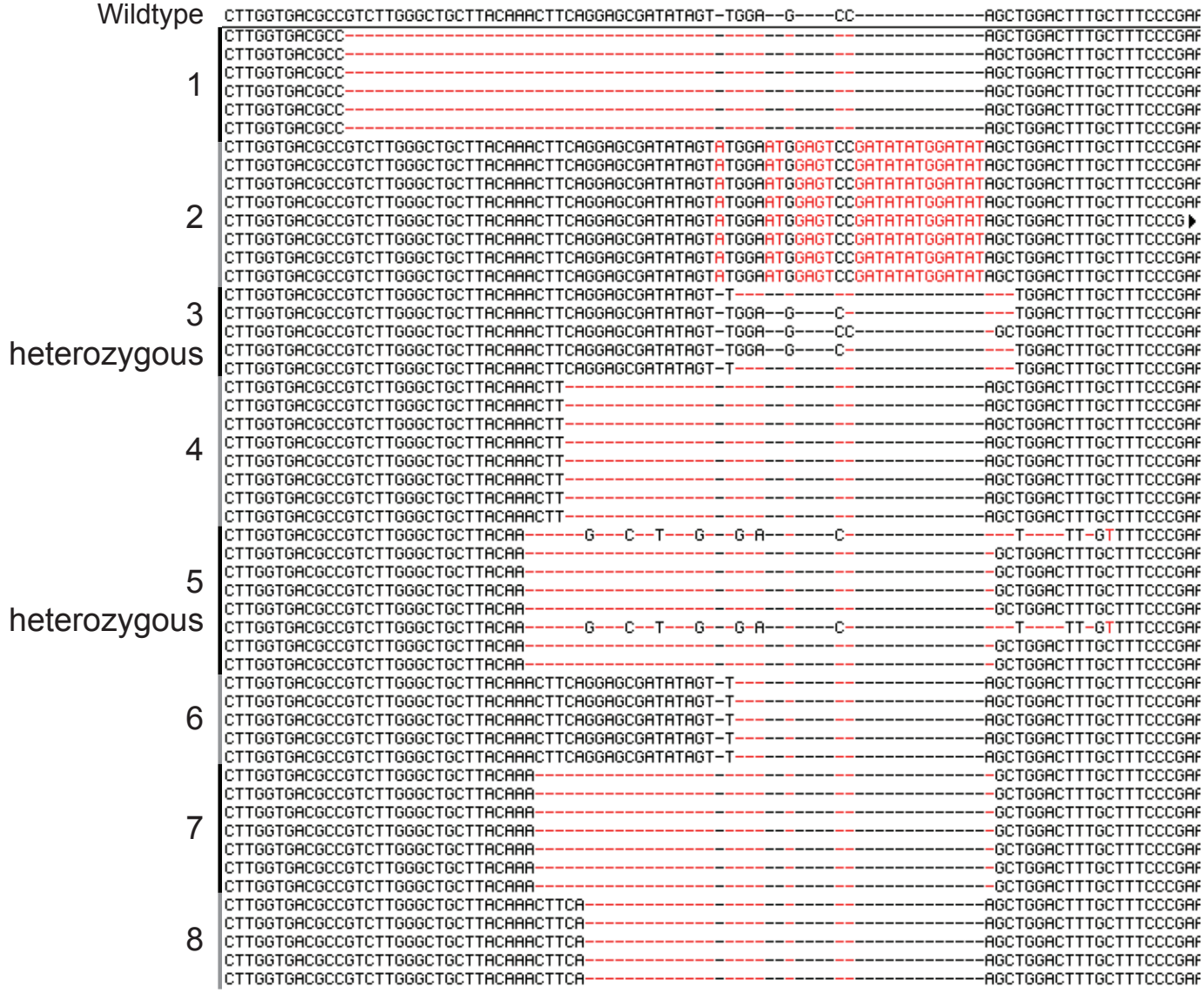
A

Media		Clones	% survival
Schneider's	10% FBS	0	0.0
	15% FBS	0	0.0
	20% FBS	0	0.0
Conditioned Schneider's	10% FBS	30	15.8
	15% FBS	31	16.3
	20% FBS	32	16.8

B



C



D

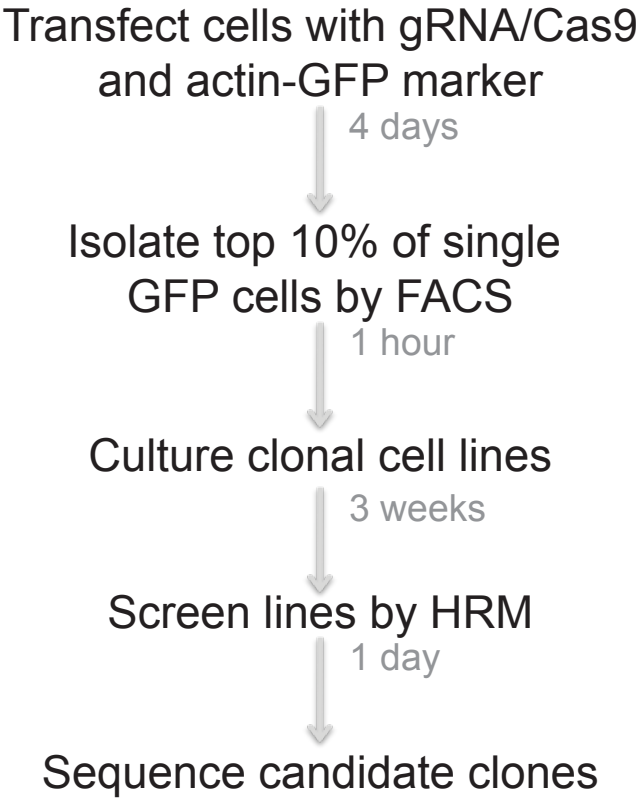


Figure_3: Generation and characterization of mutant cell lines

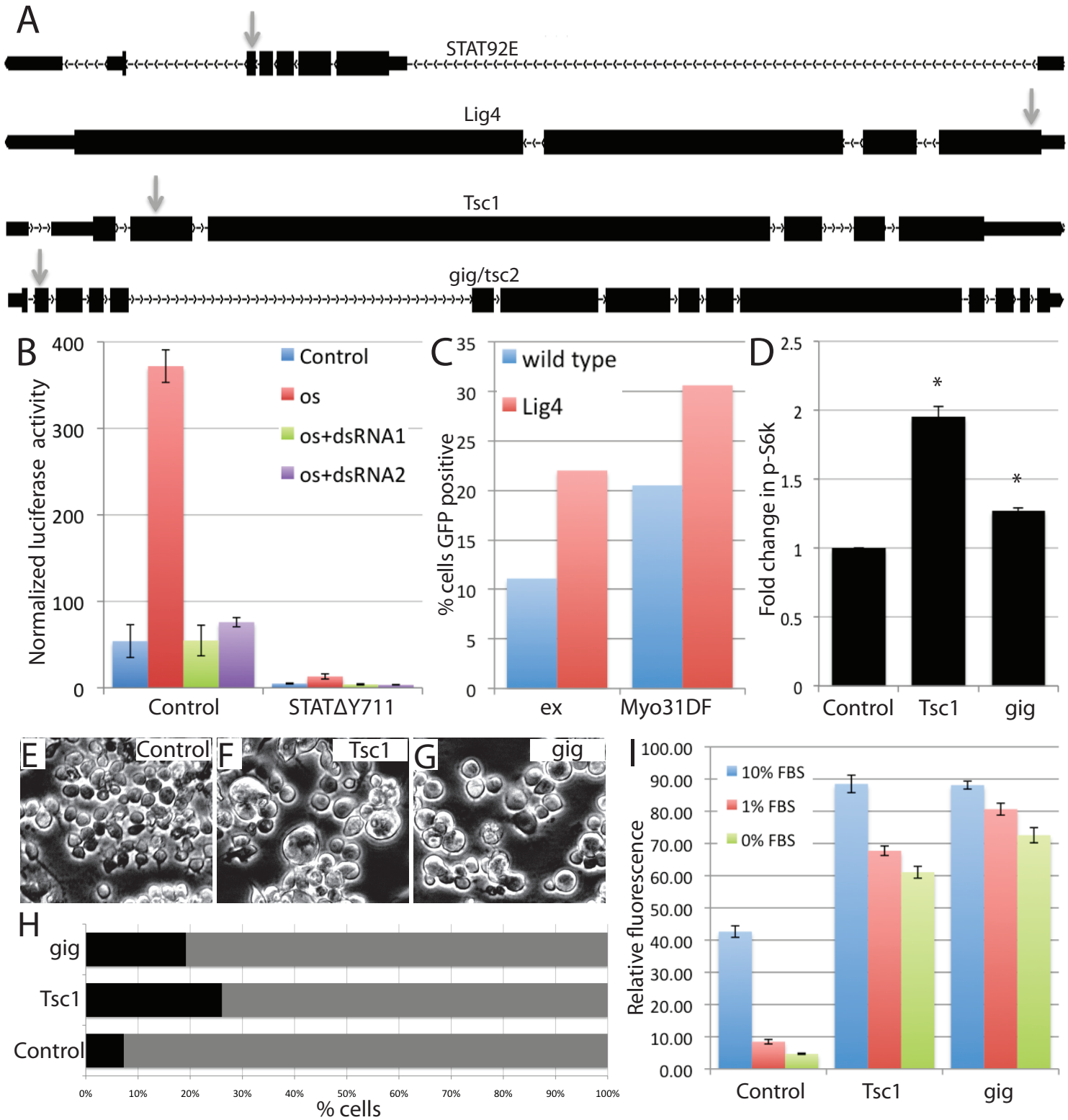
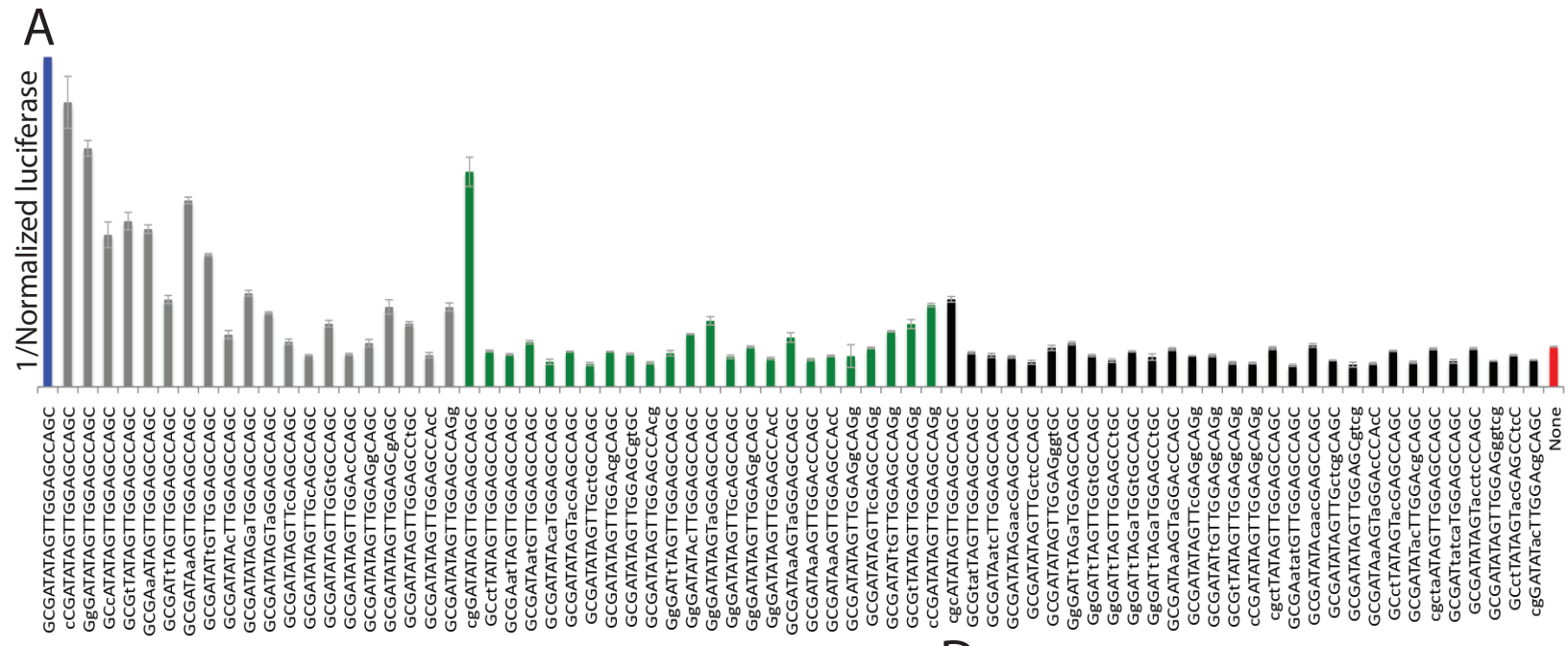
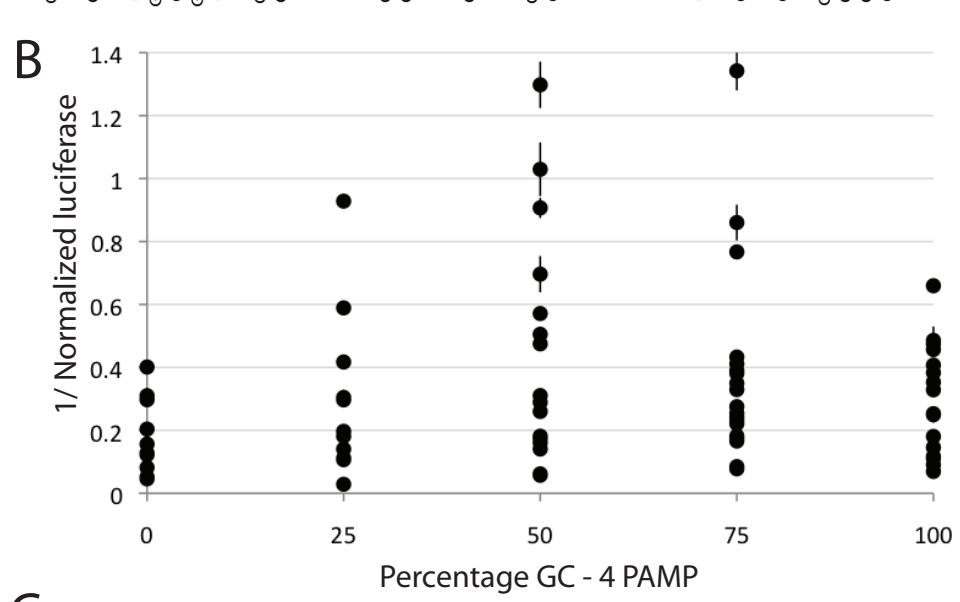
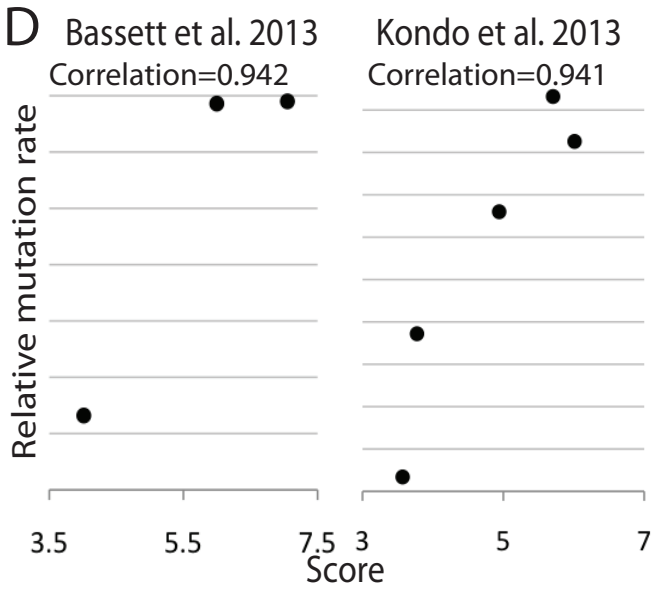


Figure_4: Optimization of the CRISPR system







C

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0.4979	0.7959	0.7553	0.6569	**0.9481**	0.7147	0.4370	0.6212	**0.9077**	**1.0000**	**0.1957**	0.7959	0.6212	0.8912	**1.0000**	0.5485	**0.9942**	0.5485	0.4550	**1.0000**
T	0.6699	0.5485	0.2750	0.5972	0.6212	0.7555	**1.0000**	0.5131	**0.8608**	0.7553	0.6569	0.3417	**1.0000**	**0.0160**	**0.9146**	0.7555	0.2906	0.4979	0.5485	0.5131
C	0.4979	0.6869	0.8528	0.7643	0.5325	0.3417	0.3417	0.7643	0.6434	**0.0092**	**0.9331**	0.5325	0.7272	0.9708	0.2905	0.7272	0.2957	0.7918	0.6434	0.5062
G	0.7918	0.4461	0.4851	0.4461	0.3417	0.6869	0.2417	0.5485	**0.0947**	**0.9256**	0.5325	0.8308	**0.1255**	0.7918	0.2544	0.4461	0.4979	0.6212	0.7918	0.4461